DOCUMENT RESUME

ED 209 711 CS 503 666

AUTHOR Goldman, Ronald J.; LaRose, Robert
TITLE Assessment of Audience Feedback Systems for Research and Programming.
INSTITUTION Corporation for Public Broadcasting, Washington, D.C. Office of Communication Research.
REPORT NO ISBN-0-89776-069-7
PUB DATE 81
NOTE 86p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Audiences; *Communications; *Feedback; Programing (Broadcast); *Public Television; Research Tools; *Technological Advancement; Television Viewing
IDENTIFIERS *Audience Response; Interactive Systems

ABSTRACT

Defining audience feedback systems as technical configurations that allow viewers in their own homes to respond electronically to television programing via telecommunications networks, this report examines research into the use of audience feedback systems as a means of reducing the cost and labor of eliciting viewer opinions and attitudes to a level that local public television stations can afford. The first section of the report specifies four applications in which audience feedback systems could benefit public television stations: pretesting, qualitative ratings, public opinion polling, and interactive programing. The second section defines these applications, describes them in terms of their uses and operational limitations, and offers examples of how each application could be implemented. The third section of the report covers the cost, limitations, and appropriateness of each of the possible applications, while the fourth section discusses the general barriers to the use of the systems. The fifth section compares each system to conventional feedback methods and to its competitors, and the sixth section notes the best technologies suited for each application. The systems are listed in an appendix. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED209711

ASSESSMENT OF AUDIENCE FEEDBACK SYSTEMS FOR RESEARCH AND PROGRAMMING

Ronald J. Goldman, Ph.D.
Robert LaRose, Ph.D.
Applied Communications Networks, Inc.
Santa Monica, California

March 31, 1981

A Report Prepared for the
Corporation for Public Broadcasting
Office of Communication Research

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Corporation for Public Broadcasting

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TABLE OF CONTENTS

PREFACE		v
I.	INTRODUCTION	1
	A. The Problem	1
	B. Approach	3
	C. Overview of This Report	4
II.	APPLICATIONS	7
	A. Pretesting	7
	B. Qualitative Ratings	12
	C. Public Opinion Polling	15
	D. Interactive Programming	18
III.	AUDIENCE FEEDBACK SYSTEMS	21
	A. Two-Way Cable	23
	B. Telephone-Based Single Response Systems	27
	C. Telephone-Based Multiple Response Systems	35
IV.	A PILOT STUDY OF COMPUTERIZED PHONE-BASED DATA ACQUISITION	51
V.	THE COST OF AUDIENCE FEEDBACK SYSTEMS	59
	A. Cost Comparisons for Pretesting	60
	B. Cost Comparisons for Qualitative Ratings	66
	C. Public Opinion Polling Cost Considerations	69
	D. Cost Comparisons for Interactive Programming	70
VI.	CONCLUSIONS AND RECOMMENDATIONS	75
	A. Conclusions	75
	B. Recommendations	79
REFERENCES		81
APPENDIX		
	AUDIENCE FEEDBACK SYSTEMS	

PREFACE

Within the emerging "new technologies" environment, foresighted planning of optimal roles for public television requires careful surveillance of the telecommunications horizon. New communication channels and new content possibilities within those channels are already presenting our industry with exciting challenges and opportunities.

The following paper was commissioned by the Office of Communication Research to survey the current and near-future telecommunications universe and to suggest potential applications of new developments both for audience research and program production. The result of the authors' overview is an intriguing menu of new audience feedback possibilities. It is our hope that the dissemination of these findings will encourage innovative approaches in the development and evaluation of new programming material.

Howard A. Myrick, Ph.D.
Director
Office of Communication Research

Carol Keegan, Ph.D.
Associate Director
Office of Communication Research

I. INTRODUCTION

A. The Problem

Many of the "blue sky" interactive residential services popularized during the early 1970s are now becoming a reality through the marriage of computer and telecommunications technologies. A variety of information and transaction services are already available which individuals can access in their homes by connecting to computer-based systems via cable or telephone networks. These services include shopping, banking, information retrieval and entertainment.

For television broadcasters, these very same technologies hold the potential for dramatically changing the traditional models of programming and audience research. They offer channels through which viewers can respond to television programs and convey their opinions and attitudes to the broadcaster almost instantaneously, thus allowing viewers to "talk back to their TV sets." Audience feedback systems will enable viewers to express their opinions about the programs they view and even allow them to be active participants in the program itself. Such capabilities may even alter the basic nature of television viewing, transforming it from a passive to an active experience.

Electronic audience feedback systems have much to offer public television stations. These technologies can give local stations tools by which they can become more responsive to the needs and interests of specialized audiences and continue to develop innovative programming concepts. This kind of capability is

especially needed at a time when public television is being confronted by competition from commercial interests who are entering into programming areas which were previously the sole purview of public television. The Office of Communication Research at the Corporation for Public Broadcasting (CPB) commissioned this paper as part of its continuing commitment to exploring innovative research methods of use to public broadcasting. Its purpose is to identify promising technologies that could meet the special needs of public broadcasters. In particular, we will explore the possibility that audience feedback systems could reduce the cost and labor of eliciting viewer opinions and attitudes to a level which could be readily afforded by local public television entities.

B. Approach

Audience feedback systems are defined in this report as technical configurations which allow viewers to respond electronically to television programming via telecommunications networks. The report is limited to those systems which are, or are soon to be, marketed by commercial venders and are accessible for use by any public television entity. In addition, the systems must allow viewers to respond to the video materials from the natural viewing environment, in the home.

Audience feedback systems were identified by reviewing communications and marketing research publications and by contacting research practitioners in those industries. Announcements were placed in leading trade journals to solicit information about relevant systems. The marketers of the various identified systems were then contacted and interviewed in person or via telephone.

C. Overview of This Report

This report begins by specifying four applications in which audience feedback systems could benefit public television organizations. They are pretesting, qualitative ratings, public opinion polling and interactive programming. These applications were formulated based upon two conferences sponsored by the Corporation for Public Broadcasting (CPB, 1980a; 1980b) and the authors' own experiences working for local public television stations. These applications are defined in Section II of this report and are described in terms of their uses and operational parameters. Prototypical examples of how each application could be implemented with an electronic audience feedback system are also offered.

The operational description of the systems and their cost, limitations, and appropriateness for each of the identified applications are covered in Section III. Some of the general barriers to utilizing audience feedback systems are discussed in Section IV based upon a demonstration of one of the identified systems. Each of the recommended systems is analyzed in Section V against conventional methods and against its competitors. Particular attention is paid to the relative cost

advantages of the different systems. Based upon the analysis of the operational characteristics, we identify the technologies best suited for each application in Section VI.

II. APPLICATIONS

A. Pretesting

Definition. Pretesting is research conducted to determine the potential audience response prior to the broadcast of a television program. Such formative research assists decision makers in approaching questions of conceptualization, development, scheduling, and promotion of a television program. Pretesting is usually conducted on one or more pilot programs, although storyboards, scripts, treatments, or concepts may also be tested. A closely related application is diagnostic testing of programs already on the air for the purposes of assessing changes in program content, promotion, or scheduling that could improve appeal or effectiveness.

Uses. The principal clients for this type of research are program producers who want to insure that their message is effectively conveyed to the target audience. By pretesting pilot programming with samples of potential viewers, audience response to the program (e.g., appeal, comprehension, and acquisition of desired knowledge/attitudes) and its particular attributes (e.g., cast, theme, setting, and style) can be evaluated and fed back into the production process. Pretesting is also of value to other program decision makers who face decisions relating to funding, scheduling, and promotion. The commercial sector utilizes pretesting extensively to help decide which television pilots and commercials should go on the air.

Pretesting has also been used extensively within public broadcasting in support of children's programming and, to a lesser extent, for adult programming. However, pretesting is by no means as commonly used within public broadcasting as it is in the private sector. Local public television productions in particular seldom have the benefit of pretesting. The cost of pretesting and/or the unavailability of pretesting skills at the local level are important barriers. Public broadcasters serve specialized target audiences that are poorly represented in the general audience samples recruited by commercial testing services. The recruitment of specialized audiences can make such services prohibitively expensive. Impending budget cuts in tax-supported programs for children (e.g., ESAA and other Department of Education programs) portend an era of less, not more, pretesting in public broadcasting generally.

This is unfortunate since pretesting could be more fully utilized within public television to assist the CPB Program Fund Director, potential underwriters, and program consortia in determining which programs should be supported. Such findings could also be useful to local stations when they plan programs, schedules and promotion strategies, or when they present their "case" for local productions at the program cooperative meetings.

Parameters. Pretesting or diagnostic testing require the collection of a relatively large quantity of data pertaining to a single program. Continuous ratings of appeal are often collected during viewing in order to determine viewer response to each segment and thereby identify the successful components

of the show. A variety of program analyzer techniques and observational measures are currently used for this purpose. Rather lengthy questionnaires are also used to obtain the social and viewing characteristics of viewers, overall ratings of appeal, evaluations of specific program components (e.g., characters, themes, and settings) and other relevant dimensions (e.g., viewing intentions, image analyses). Open-ended responses to program content may also be elicited through questionnaires or focus group sessions, either in conjunction with the closed-ended questions or in lieu of them. Public broadcasting programs impose additional data collection requirements when audience outcome objectives, as well as program appeal, must be measured.

Quick turnaround is extremely important owing to the necessity of keeping up with tight production schedules. Preliminary data must sometimes be available overnight and final results must be completed within a few days at most.

A serious problem in pretesting is how to get the test program to the respondents. The conventional method is to invite respondents to attend a showing at a central viewing auditorium, e.g., Audience Studies, Inc.'s Preview House. However, this arrangement has little external validity and imposes a considerable burden on respondents, especially in an era of rising transportation costs. In order to test programming in a more natural environment, some commercial organizations (e.g., Audience Studies, Inc., Burke Marketing Services, and Blumenthal Research Strategies, Inc.) utilize cable systems for transmitting

pilots, and viewers' responses are collected through diaries and/or subsequent telephone interviews. Public broadcasters do have the option of broadcasting a pilot, a practice which is an unacceptable security risk in the highly competitive world of commercial broadcasting. However, the logistics of data collection from the home using conventional methods lengthen the turnaround time considerably -- or increase the cost.

In order to meet these essential requirements at an affordable cost, sacrifices are often made regarding the quality and size of the sample. It is common practice to utilize quota samples of a target audience, usually not selected on a purely random basis, and sample sizes of only a couple hundred. Response rates of only 10 or 20 percent are accepted in order to reduce the lead times and costs associated with recruiting. Tests are often conducted in one or two of the major markets and thus fail to represent truly the audiences of local broadcasters.

To compensate for the nonrandomness of the sample, research organizations resort to the use of norms which tell them how a program compares to other similar programs rated by audiences recruited by identical (if nonrandom) means. Pretesting for public broadcasting poses a special problem in this respect since highly segmented and varying target audiences, rather than the general viewing audience, are of interest. This factor can add greatly to the cost of research and frequently invalidates the norms that are the key to interpreting the data gathered by commercial research services.

Prototype Example. A prototypical pretest involving interactive technologies might entail collecting program analyzer and questionnaire data from 400 people in a single local market. We will assume that the program is 30 minutes long and that continuous ratings are made every minute. The salient attributes of the program would be rated after viewing, and the audience's comprehension of the program's salient points would be assessed. Open-ended responses would also be elicited from 20 percent of the sample who had the strongest reactions (positive or negative) to the program. Basic demographic and viewing characteristics would also be specified. A total of 60 columns of data would be obtained from each subject. The sample would be recruited on a quota sampling basis from an expanded telephone frame to include PBS viewers and representatives of special target audiences. Respondents would be invited to view the test program in their homes via open-broadcast. Initial results would be available to the local station that commissioned it on an overnight basis with an in-depth report completed within ten working days.

B. Qualitative Ratings

Definition. Most on-going evaluations of television programming are quantitative in nature, stressing the measurement of audience size, e.g., the Nielsen ratings. Such measurements serve the purposes of the commercial broadcasters whose objectives are to attract the largest possible audience for their advertising clients. Quantitative ratings are the primary datum of the broadcast industry. They are used to frame the full range of decisions involving programs, scheduling, and promotion. On the other hand, public television serves many small and specialized groups which exist within a community whose interests are not otherwise met. In this perspective, public television should not be judged solely on the basis of audience size, but by the degree to which these specialized viewers and their interests are served. Public broadcasters are more concerned with the public's qualitative ratings of programming. That is, they need to know how specialized audiences rate programs on dimensions such as enjoyment, originality, community service, personal relevance, etc.

Uses. The concept of qualitative ratings and their importance to public television has been extensively discussed in previous reports (CPB 1980a; 1980b). Briefly, it is intended that qualitative ratings will be used by public broadcasting executives at the national and local levels to make decisions about the funding, scheduling, and promotion of PBS programs. It is also hoped that the system will provide an additional dimension of accountability to both private and public sector underwriters

of public broadcasting. Crude versions of qualitative ratings (e.g., TvQ) are already used by commercial broadcasters.

Parameters. In contrast to pretesting, the quality of the sample is paramount for qualitative ratings. Since the results are to be generalized to all television viewers, it is important that the sample be truly representative of the national audience of PBS viewers and randomly chosen. A high degree of quality control must be maintained. Respondents must be identifiable and verifiable, and every effort must be made to obtain high (50-60 percent) return rates. Moreover, an extremely large sample (i.e., several thousand) is required. Since public television attracts relatively small audiences and is addressed to the needs of highly specialized segments, the sample must be large to ensure that each program has enough respondents from each important audience segment to produce a valid rating for each. Participants provide a relatively small number of responses per program (i.e., rating each show on four or five scales), yet a high volume of data would be generated, since many programs would be rated each day by a large number of respondents. Such a system would operate on a continuing basis, perhaps with several "sweeps" over the course of a year. It might be operated on a panel basis, with each respondent staying in the sample for more than one sweep. Qualitative ratings do not require fast turnaround of results, with a period of days or weeks acceptable.

Prototype Example. The application, as envisioned in this report, would consist of an on-going system by which samples of

television viewers would evaluate the quality (as measured on four or five specific dimensions) of the various programs (both commercial and public) they view. Qualitative ratings would be collected from a national stratified sample of 3,000 households, screened for regular PBS viewership and for membership in important demographic (e.g., minorities) and interest (e.g., opera lovers) groups. All members of a household would be asked to participate, and each household would stay in the sample for four sweep periods over the course of the year, generating an average of 6,000 columns of data per household per year. Preliminary qualitative rating reports would be available within one week of each sweep, with final results available inside of three weeks.

C. Public Opinion Polling

Definition. Polls of public opinion are conducted to ascertain the state of public attitudes toward various issues and political candidates among the general population.

Uses. Public opinion polling could be used by public television stations to identify community needs and to facilitate the discussion of public issues. One application would be an extension of existing community ascertainment procedures on which program development and policy decisions could be based. Another case is a program-related function in which the opinions of a representative sample of the community could be used in public affairs programming, similar to the telephone "insta polls" which many major market commercial stations routinely report as part of their news broadcasts, but conducted while the program is in progress.

Parameters. Similar to qualitative ratings, the results of public opinion polling are to be generalized to a larger population. Very high standards must be met for sample selection and quality control. Biases arising from panel studies which might be acceptable for qualitative ratings do not meet the high standards set by public opinion pollsters. A large sample is needed, though not as large as for qualitative ratings since the opinions of the general public, more than highly specialized subpopulations, are of interest. Sample sizes of 400-1,100 are suitable, depending on the amount of sampling error acceptable.

The use of polling in conjunction with a television broadcast

will pose some peculiar problems not usually encountered in public opinion research. The system will have to have a very high peak load capacity in order to accommodate a large number of respondents accessing it within a short time frame. Controls will have to be exercised to insure that only preselected, representative respondents are allowed to register their opinions, and that each responds only once. Otherwise, the results will be useless as a tool for public debate. Though dependent on the specific application, quick turnaround of data analysis is of high importance. If the poll is to occur during a televised program, results will be needed within minutes, or at least by the time of the next regularly scheduled broadcast. The system utilized must also be able to provide anonymity in order to protect the respondents' identity when asking sensitive questions about important social issues.

Compared to qualitative ratings, such polls would require the collection of a relatively small amount of data from each respondent. The system must be more flexible so that questions could be continually changed. Respondents would remain in a sample for a limited time, perhaps only for a single study.

Prototype Example. In one application which inspired this study, a local station wanted to participate in its community's goals conference. The program concept (which was not implemented) was to invite the city fathers to present and debate the various issues facing the community on a television program, followed by a poll of a representative sample of the community to provide immediate feedback while the program was still in progress.

Within a community, 400 potential respondents were to be recruited in advance from an expanded telephone frame and agree to view the program. The sample was to provide answers to a few questions during the course of the program at particular points when a question was posed. Results were to be available within ten minutes to maintain the flow of the program.

D. Interactive Programming

Definition. The growing availability of audience feedback technologies offers a variety of intriguing "interactive television" possibilities which will allow the responses of viewers to become an integral part of the programs that they watch. In essence, this application transcends the traditional nature of television by transforming viewing from a passive to an active experience. It differs from the previous (public opinion polling) application in that the population of interest is viewers of a particular program, rather than the community at large.

Uses. Possible scenarios for public broadcasters would include having the audience direct improvisational drama, take part in televised town meetings, take an active part in live interview shows, vote for musical selections they would like to hear, make bids at station auctions, or provide feedback to instructional television lecturers. The limitations of using such a capability are only that of the producer's imagination. Presently, the only means by which such participatory programming can take place is by having viewers call in, but only a few individuals can have input into any one program.

In such applications, the goal is not to measure public opinion reliably and validly as much as it is to stimulate viewer involvement. Limited experiments in this area indicate that allowing viewers to interact with a program does produce increases in audience shares (Media Science Newsletter, 1980).

Parameters. The essential characteristics of the interactive programming application are quick turnaround and the ability to handle an exceptionally high peak load capacity as large numbers of viewers try to call in at the same time. There is no ideal sample size. Rather, the principle would be "the more the merrier." In fact, the number of people calling regularly could become a continuing measure of the program's popularity. The representativeness of the sample and quality control requirements are unimportant. The group which participates (regardless of who they are) is the population of interest.

Prototype Example. A local PBS station in a major market has a weekly call-in feature (e.g., "Your town meeting of the air") as part of a local public affairs program. A single question would be posed early on in the program which could be answered by picking one of two to five alternatives (e.g., "How much should the city council spend on police protection? a. $1 million b. $2 million c. $3 million d. $4 million e. over $10 million). Viewers would then have 20 minutes to phone in their answers. Results would be presented before the end of the program and discussed by commentators or panelists. On any given week approximately 4,000 calls would be placed. To prevent the lines from becoming overloaded, potential respondents would be preselected. They would send in stamped, self-addressed post cards along with their request to be included in the town meeting panel. Cards would be randomly drawn and the selected respondents would be notified of the call-in number and their own call-in

time (perhaps scheduled by five-minute periods after the announcement of a question) by return mail. Whenever the number of callers dropped below 4,000 for a particular week, new callers would be introduced into the sample. To give the largest possible number of viewers a chance to enter the panel, the call-in numbers could be changed periodically and a new set of panelists drawn.

III. AUDIENCE FEEDBACK SYSTEMS

This section provides an overview of a variety of technological systems which could be utilized by local public television organizations for the previously discussed applications. Inclusion of potential audience feedback systems have been restricted to those systems which are now, or soon to be, marketed for research purposes and available to public broadcasting organizations.

The systems have been categorized according to two dimensions. The first dimension relates to the network configuration upon which the technologies are based. There are two principal network technologies which are applicable to this assessment of audience feedback systems: cable and telephone networks. Cable is a broadband technology which is utilized to distribute television programming to residential subscribers. When equipped with the appropriate hardware, cable could allow narrowband, digital information to be transmitted from homes to a central computer located at the cable system's headend. The telephone system, as it has existed for some time, is an interactive, narrowband medium by which voice and digital (Touch-tone) signals can be transmitted from any location to any other location on the network. Radio-based systems could be conceptualized, but in practicality are not appropriate to the applications at hand.

A second dimension pertains to the amount of data which the system can accept in a single transaction. This dimension is particularly important in discriminating between the various telephone-based systems. One set of technologies is single-response systems in which only a single, discrete response to a single, closed-ended question is possible. A second set of systems will accept multiple responses in a single transaction.

Each individual technology will be discussed in terms of operational description and costs. Because of the similarity among technologies within each category, an evaluation of the systems' limitations and their appropriateness to the identified applications will be discussed at the end of each subsection.

A. Two-Way Cable (Broadband Multiple Response Systems)

1. Systems

a. QUBE. Since the early 1970s, Cable has been regarded as the revolutionary means of providing interactive services to the home. In fact, it was the Warner/Amex's QUBE cable system in Columbus, Ohio, that brought "interactive television" into reality. Viewers could "talk back to their sets" by pushing one of five buttons on their cable converter (which contains a microprocessor), and their responses were continually polled by the central computer at the cable headend. A question could be asked and the results tabulated within 20 seconds.

The QUBE system gained notoriety for being able to conduct informal polls of the viewers and to provide interactive television. On one occasion, QUBE subscribers were allowed to vote on questions raised by the Commissioner of the U.S. Food and Drug Administration. NBC used QUBE to elicit responses to a presidential speech, while another time, viewers had an opportunity to call plays for a local junior college football team. In an experimental research application, CPB used QUBE to identify evaluative words and phrases viewers use to distinguish more and less preferred television programs.

b. Other Systems. Since the time the Columbus QUBE system was put into operation, the availability of new hardware configurations has greatly enhanced the potential of two-way cable systems.

For example, QUBE III, manufactured by Pioneer, allows home converter/terminals to communicate with the central computer at shorter intervals using up to eight numerical digits of variable length data (i.e., responses do not have to be in a set format). In addition, each converter is addressable: the computer can identify responses from each household and maintains a high level of redundancy for correcting errors. Similar technology is being developed by Tocom, Oak, Jerrold, and Scientific-Atlanta.

2. Limitations of Broadband Multiple Response Systems

The main drawbacks to two-way cable technology lie not in the technology itself, but in its limited availability. Although QUBE is over three years old, Columbus still remains the only location in which two-way cable capabilities exist. In fact, the QUBE system is available in only one section of that city. Although many franchise applicants are now proposing the speedy deployment of two-way services, the necessary hardware and software developments are seriously lagging, raising doubts about the ability to deliver on these stated promises (Cablevision, 1981).

In the near future, fully two-way cable services are likely to be available in only a few scattered locales. Warner/Amex is building QUBE systems in Cincinnati, Pittsburgh, Dallas, Houston, and in a number of smaller municipalities, but most of these systems are several years away from completion. Cox Cable is also leading the way in developing two-way services with its INDAX system. INDAX

is presently being implemented on the Mission Cable System in San Diego, California (the largest cable system in the country) and is promised for recently awarded franchises in Omaha and New Orleans. Thus, the adoption of two-way cable will proceed very slowly. It will be decades before it (or any form of cable television) is used by a cross section of the general population. Even where available, two-way cable systems will not be interconnected and compatible, since each municipality issues separate cable franchises to different operators. Thus, aggregation of data across geographical areas may be difficult.

How two-way cable will be utilized, even when available, is not totally clear. The most likely uses of the two-way capability will be in those areas of the highest economic payoff, e.g., security, banking, shopping, cabletext, energy management, and pay-per-view programming. Although the use of two-way cable for research purposes has been cited as a possible application, few cable operators have begun gearing up to offer such a service beyond one-shot, special studies. To support an on-going research service, appropriate computer software must be designed, specialized personnel employed, and additional hardware added to interface with the central computer. Concerns have already been raised about potential threats to privacy posed by the integrated data bases which could be generated by two-way systems (Los Angeles Times, 1981). For the most part, it must be remembered that cable operators are primarily in the cable television business, and research services might seem like more trouble than they are

worth to cable operators.

3. Applications of Broadband Multiple Response Systems

Technically, two-way cable systems are excellent for all suggested applications. The direct connection of the viewer to a centralized computer allows rapid and accurate data collection in both continuous and discrete modes. Turnaround time is quick, peak capacity is unequalled by narrowband systems, and large sample sizes can be easily managed. The ability to address specific households makes each respondent identifiable (for quality control purposes) and allows test shows to be narrowcast to only those participating in a particular study.

While two-way cable meets the primary requirements of the pretesting interactive programming applications in markets with two-way systems, it has limited applicability in qualitative rating and public opinion studies. The major limitation is that not all the residents of a given geographical area may be two-way cable subscribers, which would not allow a representative sample to be drawn. In addition, the incompatibility and uneven distribution of the various cable systems across the country will not allow truly national samples to be drawn for any purpose for decades to come.

4. Costs

Since research services on two-way cable systems are not currently available on a continuing operational basis, it is difficult to

assign a cost to them. It is not appropriate to use the custom studies conducted by QUBE as benchmarks because they were "special" and do not truly reflect the operational costs which would be incurred on a regular basis. The cost of each individual terminal is about $230. The terminals are usually offered as part of a "tier" in cable system price structures at an additional $3 per month over and above the basic cable service. In addition to the terminal rental fee, subscribers pay a monthly fee to use the service. Typical costs for a simple two-way service (e.g., security or fire protection) are $20 per month per subscriber. Using this as a base cost, if a research organization wished to introduce a respondent into the two-way cable network for purposes of data collection, it would cost the researchers at least $23 per household per month. The costs of recruiting subjects, data analysis and narrowcasting test material to respondents would be added to this figure.

B. Telephone-Based Single Response Systems

1. Systems

a. Dial-It

1) Description. The American Telephone & Telegraph Company (AT&T) began offering a mass calling service in September 1980 called Dial-It. The service is designed to serve mass calling applications. The application of interest here is Media Stimulated Calling in which a large number of callers attempt to reach a single set of telephone numbers as a result of advertising the

number on radio or television. Dial-It channels the calls into special nodes which can handle heavy traffic, thus preventing interference within the regular switched network facilities. The Dial-It system is comprised of seven nodes, each of which has a capacity of serving as many as 1,800 callers simultaneously. Since the system operates within the structure of the telephone network, multiple telephone lines to private locations are not required, and the system can be accessed by any telephone.

Dial-It can be utilized as an audience feedback system by asking callers to express their opinions on one question at a time by dialing one of the several telephone numbers, each associated with a specific answer to a closed-ended, multiple-choice question. A possible scenario would be the presentation of a question during a television program in which viewers are instructed to dial one telephone number for answer A, another telephone number for answer B, etc. The results of the poll would be determined by the number of people who call each of the respective telephone numbers in a given time period. Daily reports of the results are provided as part of the basic service, and minute-by-minute tallies of calls are available for an extra hourly charge.

This system was utilized by ABC to poll viewers of the Reagan-Carter debate in October 1980. Almost 725,000 people participated during the 100-minute period following the debate, with 477,815 callers dialing the Reagan number and 243,554 dialing the Carter number.

Dial-It also allows randomly selected calls to be forwarded to specific telephone numbers, whereupon the sponsor can directly communicate with the caller for in-depth interviews. Otherwise, the callers simply receive a prerecorded message thanking them for participating in the vote. Technically possible, but not currently offered by AT&T is voice storage, in which the caller can leave a verbal message or a string of digital impulses input from Touch-tone telephone keypads.

2) Costs. The cost for using the Dial-It service includes charges to the sponsor and the caller. There is a $5,000 annual fee charged to sponsors for access to the service. In addition, there is a charge of $500 per announcement, per day, per time zone for the basic service package, which includes (1) cut-through capability, or the ability of the sponsor to interview selected callers; (2) a 15-second announcement capacity; (3) multiple telephone numbers which are assigned to closed-ended responses; (4) announcement updating as required; and (5) daily volume counts. Additional charges apply to longer announcements (i.e., $100 for a 25-second announcement) and minute-by-minute tallies ($50 per hour). These charges are not day-segment sensitive, but the telephone company reserves the right to limit which times the service is available.

Equally important is the charge to the caller for utilizing the service. Fifty cents per call will be charged to each user, which will automatically be added to their monthly telephone bill. In the Reagan-Carter application, AT&T not only received

$7,000 from ABC for the use of the service (including the annual fee), but also over $360,000 in charges collected from callers.

b. VOTRAK

1) Description. VOTRAK, marketed by Unlimited Television, Inc. (UTI) of New York, is promoted as a "two-way" television system. It operates much like the "Media Stimulated Calling" of AT&T's Dial-It service. In fact, VOTRAK's inventor, John J. Root, claims that AT&T has infringed upon his patent and plans legal action.

Similar to Dial-It, VOTRAK allows television viewers to respond in real time to a closed-ended, multiple-choice question. Each designated answer is assigned a unique telephone number which the viewer dials on his/her telephone in order to "vote" for that particular alternative. When the connection is completed, callers hear a short recorded message thanking them for participating and telling them that their "vote" has been registered, and then they are disconnected. Results of the number of calls received by each telephone number are then tabulated. A television camera pointed at counters on the side of the device could allow the results to be superimposed on the television picture during the voting. In addition, the system enables calls to be intercepted for further questioning of the respondent by an interviewer or even by an on-air host.

Since all the system does is count calls, any telephone could access the system, but only one question can be voted upon in

any one call. Connect time with the system is approximately six seconds, thus allowing about ten callers to be processed per minute per telephone line.

Multiple telephone lines are required for each answer in order to handle a high volume of calls within a short time period. The number of telephone numbers required is a function of the size of the viewing audience, the percent of viewers who participate, and the duration of time in which voting is allowed. UTI states that 20 telephone lines would be sufficient to handle responses for an hour television program with a "2" rating in New York.

VOTRAK was utilized by a local New York television station (WNEW-TV) in August-September 1976, as part of its Ten O'Clock News program and on the New York City cable systems (Manhattan Cable/Teleprompter) in January-April 1978. In the first application, viewers were polled as to their opinions of the "question of the day" with "yes" and "no" responses possible. The number of respondents varied according to the question asked, with participation ranging from 4,150 to 25,997 and averaging at 12,470. UTI calculated that an average of 2.9% percent of the average quarter-hour ADI homes reached by the program responded and that viewership of the program increased 40 percent during the period of the VOTRAK test.

2) Costs. VOTRAK is available on a monthly lease basis. UTI quoted CPB a rate of $6,000 per month for a 20 telephone line capacity and for up to ten uses. There are additional charges

of $600 for each time above ten, an installation fee and telephone line charge to be paid to the telephone company. The latter might be absorbed in overhead by stations which maintain the multiple lines used during pledge and auction periods. It must be noted that these charges are probably negotiable, since they fluctuate across UTI documents.

c. Telephone Answering Machines

1) Description. Dial-It and VOTRAK operate as telephone answering devices, which simply play a prerecorded message and count calls. There is no reason why such a system could not be assembled by purchasing several telephone answering machines, attaching them to telephone lines, and assigning their telephone numbers to different response alternatives. Results would be determined by summing the number of calls recorded on the appropriate counters.

2) Costs. A configuration of telephone answering devices could be assembled for an up-front capital cost of approximately $150 per telephone line. The only operating expenses would be the cost of the multiple telephone lines, which many public stations maintain anyway, and the cost of personnel to tally the calls. However, answering machines would work more slowly than VOTRAK so that a larger number of lines would be needed. The volume VOTRAK could handle with 20 lines would probably require 75 lines with answering machines, an initial outlay of over $10,000.

2. Limitations of Telephone-Based Single Response Systems

The major weakness of these systems is that they cannot handle multiple responses, i.e., only one response can be recorded per interaction with the system. Without the ability to electronically transmit individual identification codes, respondents cannot be readily identified. This maintains anonymity but does not allow the kind of quality control essential for public opinion polling or qualitative ratings.

In addition, the fee that Dial-It passes on to the caller could lead to "class bias," in utilizing the system. Lower income individuals would be less likely to participate if they are required to pay a fee. Charges of "elitism" were levied against ABC by pollsters following its Dial-It application in the Reagan-Carter debate.

Another problem is that even very large numbers of telephone lines cannot handle the peak loads generated by media-stimulated calling. Although 320 telephone lines were utilized in VOTRAK's WNEW trial (of which half the lines were for a positive response and the other half were for a negative response), only 46.8 percent of the attempted calls were completed during the hour program according to a New York Telephone company memo. With such a low completion rate for phone calls and with the self-selection sampling bias, the results of the poll cannot be said to reflect the opinions of poll participants or of the WNEW viewership. For example, the true distribution of answers to a question could be 70 percent-30 percent, but if only an equal number of

people can get through for each side, the results would be 50 percent-50 percent. National Public Radio reported that a similar problem arose during its Dial-It application, when callers from urban areas (presumably Democrats) had more difficulty in getting through than callers in rural areas (presumably Republicans).

With possible exceptions in smaller markets, none of these systems can be counted on to provide equal access to all potential callers. Therefore, some kind of preselection process (e.g., viewers apply in advance by sending in stamped, self-addressed post cards) is in order.

Finally, the self-selection bias also prevents generalization of the results to the community. This flaw was emphasized following the New York cable test and resulted in the disuse of the system on the government access channel.

3. Applications of Telephone-Based Single Response Systems

The single response capabilities of these systems immediately eliminate their consideration in pretesting and qualitative ratings applications where a large amount of data must be collected from each user. In public opinion polling and qualitative rating studies, these systems are not desirable because of the lack of quality control over the sample. Public opinion polling requires a high quality random sample, whereas these systems rely on the callers to self-select into them. Preselection of a random

sample of individuals and obtaining their cooperation in advance might help matters. However, the necessity of having to view a particular program at a particular time would probably reduce the response rates to levels unacceptable to public opinion pollsters.

The single response systems are clearly best suited to interactive programming applications in which viewers can respond to questions posed one at a time during a show. The Dial-It service is technically superior to VOTRAK by virtue of its greater peak load capacity and national reach, but the 50-cent charge to callers could limit its accessibility to a large percentage of viewers. Utilizing a bank of telephone answering machines could prove cumbersome but might perform as well as VOTRAK given enough telephone lines. It must be kept in mind that results obtained from any of these systems should be used only for entertainment purposes, and no generalization of the findings should be made to populations other than those who actually call in.

C. Telephone-Based Multiple Response Systems

1. Systems.

a. VOXBOX

1) Description. R.D. Percy Company operates a qualitative research system in Seattle, Washington, utilizing an electronic home response terminal called the VOXBOX. The Percy System includes 200 households which are asked to evaluate the television programs that they watch on an eight-point scale. The units are connected to the television set and also serve as a

channel selector. They are connected by means of dedicated telephone lines to a central computer, where qualitative responses and channel selection data are continually monitored throughout the day. This service provides syndicated reports of the qualitative evaluations of the programming and commercials in the Seattle area, much the same way that Nielsen and Arbitron provide quantative viewership data for television and radio.

2) Costs. The Percy Company sells its Television Program Preference Guide, containing summaries of qualitative responses to programs at a cost of $950 per month. Single-program results (e.g., when a pilot is shown without recruiting to view) are available for $500 per show. A pretest for which panel members would be contacted in advance and invited to view would cost $1,200.

b. PEAC

1) Description. The Program Evaluation Analysis Computer (PEAC), designed and marketed by PEAC Developments, has been extensively utilized as a formative research tool to evaluate radio, television and film material (e.g., Chen, et al., 1979). The current version of this system, PEAC I, is designed to allow an audience to respond to test materials by inputting numbers into hand-held microprocessor units. The response units are programmed before the testing period by a microcomputer. Following the session, the units are placed into a case at which time the stored data are

loaded into the microcomputer for analysis. The results are presented by means of a color graphic display, which can be superimposed over a replay of the test materials. Although PEAC I is portable and flexible, it can be used only in controlled settings, and units must be returned to their case for programming and data retrieval. A more sophisticated version of the PEAC system, PEAC II, is under development and would be placed in individual homes. The units would communicate with the microcomputer by means of telephone lines. The response unit can be wireless, and it would fit into a modem connected to a telephone jack, when not in use. Each unit would have a capacity of 500 data points and could operate by sampling responses at five- to 60-second intervals or by recording discrete answers to multiple-choice answers.

The system could operate in one of two modes. In the first, communications between the central computer and the home terminals would be fully automatic. The computer would have a list of respondents' phone numbers and would dial each in succession, giving the respondents verbal instructions for the test session and programming the response units. After the program, the computer would redial the units and collect the data. Neither researcher nor respondents would be burdened with telephoning.

In the second mode, communication between the researcher and respondent would not be fully automatic. The researcher would manually dial each respondent and give instructions for the test

session, and the respondent would call back the central computer after viewing the test program. Programming the response units and collecting data would still be done automatically by the computer.

2) Costs. The PEAC II system would be easily implemented in any household, though high unit cost of the terminal would necessitate having the units remain in the same locations over some period of time. Although PEAC II is not currently marketed, its anticipated costs are between $39,000* and $42,000 for a system including 30 in-home response terminals, central computer, necessary software, manuals, and training. Each additional in-home response terminal would cost $650, and automatic dialing of respondents and accessing of data would be another $6,000. The unit would have to be installed and disconnected by a skilled technician, and respondents would have to be trained in its use. PEAC estimates the cost of evaluating a 30-minute program with 100 in-home respondents would be approximately $3,000 with data available the next day. However, these costs apply to studies in a single metropolitan area. Installation, training and communication costs might increase dramatically for a national or multi-city sample.

c. Computafone

1) Description. Computafone, marketed for media research by Media Science Measurements of New York, offers a versatile

*In Canadian dollars

system for conducting telephone surveys. It can automatically dial telephone numbers, based on inputted telephone numbers or random numbers, ask questions on prerecorded tapes, and accept responses by means of Touch-tone impulses, rotary pulses, and verbal messages which are recorded on tape. Thus, both quantative and qualitative information can be elicited from respondents.

The unique characteristic of this technology is the ability to accept both digital and analog data. For rotary telephones, the pulses of the dial are translated into digital form. In order to correct for pulse slippage (i.e., in order to identify precisely which number is dialed), the respondent is asked to dial a specific number at the beginning and at the end of the interview, and the computer automatically corrects all responses to that basis. In cases when the two test responses are not internally consistent (which is said to happen rarely), the respondent is asked to repeat the answer verbally and the responses are recorded on tape.

2) Costs. The system will be available soon in two forms. Individual units will be sold with one to three telephone line capacity, ranging in price from $6,300 to $10,000. In addition, regional bureaus are being formed from which time can be leased. The initial bureau will be opening in New York in August-September 1981 and will focus on media research. Already, 20 advertisers have been said to have become charter members of the bureau. The cost of utilizing the bureaus will be $500 per hour for use of 45 lines, though price will vary according to the exact specifications of the services required.

d. The Communicator

1) Description. The Communicator, manufactured by Information Technology of Fresno, California, is principally designed and marketed as a telephone dial access information service. Dial access connotes systems which allow callers to dial a telephone number in order to request an audio cassette to listen to. One such service that is now offered across the country is Tel-Med, which provides medical information to the public.

Usually an operator is required to answer phone calls and to insert the requested tape into the player. The Communicator improves upon this procedure by allowing the caller to request messages by inputting the appropriate codes using a Touch-tone telephone, thus reducing personnel requirements. The system has a stored message capacity of 900 minutes which can be segmented in as many ways as desired.

The Communicator also provides a printout of the frequencies with which each message has been accessed during a specified time period. It is this feature which Information Technology has been marketing for conducting polls. The codes, instead of representing stored messages, could be assigned to answers of closed-ended, multiple-choice questions. The unit can accept only three digits per telephone call, meaning that a maximum of three questions could be asked. The questions might be posed before the caller dials up the system, e.g., during a television program or by a mail questionnaire, or in the opening message the respondent

hears after calling the service. At the end of the allotted response period, the frequency of each response can be printed out at a push of a button.

2) Costs. The unit costs $5,700, and each unit can handle only one telephone line. Although operator intervention can occur when callers do not have push-button telephones, the system would only be practical in the automated mode. To date, this application of the technology has not been used by clients for research purposes, and so no firm cost data are available for such applications.

e. Audio Response Service (ARS)

1) Description. ARS is offered by the Service Bureau Corporation of Greenwich, Connecticut. The system allows callers to access computerized information banks with their Touch-tone telephones. In a research application demonstrated by us for another project, ARS was used to collect a series of responses to closed-ended questions which users entered automatically through hand-held automatic telephone dialers (see following section). The system has a computer-synthesized voice which directs the user through the transaction. The computer could be used to ask questions, although the quality of the voice is rather poor. Data could also be input manually by owners of push-button telephones. The system is served by 18 WATS lines and can be accessed from anywhere in the country. Up to 256 digits can be recorded per call.

Researchers can access a wide variety of statistical packages also offered by SBC on the same basic computer system so that data can be turned around within minutes.

2) Costs. ARS charges are calculated by the amount of connect time at $20 per hour. The clock starts running at the moment the users identify themselves to the system with their user access code. There is a minimum charge of $100 per month for the service. Additional costs would be charged for programming changes in the user dialogue and for data processing.

f. PAL-Consumers' Computer

1) Description. Product and Area Locater (PAL), developed by San Diego-based Consumers' Computer Corporation of America, allows individuals to access computerized consumer information services, e.g., where to buy a certain product, through the use of a Touch-tone telephone. After dialing the computer and hearing a welcoming message, the caller inputs a residential ZIP code and the code number of the service desired which is identified in the PAL guidebook. The computer's voice reads back the telephone numbers of up to three businesses in the consumer's ZIP code which provide the desired service.

PAL also has the capability for conducting public opinion research. Respondents can express opinions by inputting responses using the keypad of a Touch-tone telephone based on questions asked by the computer (which can branch to appropriate questions based on the caller's responses) or prior instructions distributed via newspaper,

mailed surveys, or panel instruction packages. The current configuration of PAL has a capacity of handling only 15 local telephone lines simultaneously. There are plans to establish local PAL bureaus throughout the country. A detailed description of a research demonstration involving this system is found in the following chapter and in Appendix A.

2) Costs. PAL is not primarily a research service and so there is no firm cost structure for such purposes. A fee of $500 was charged for the demonstration project described in the following chapter. The fee covered programming, data collection and transfer. A cost of ten cents per transaction has been estimated for research applications.

g. Automatic Dialer Technology

1) Description. Useful adjuncts to telephone-based, multiple response systems are portable telephone dialers. These devices are not in themselves audience feedback systems, but could augment systems like Computafone, ARS and PAL in two important ways: (1) They give users who do not have push-button phones access to the system, and (2) they can dramatically reduce computer connect times by storing data "off-line" and playing it through the phone system in a single burst. The latter application was recently demonstrated by Applied Communications Networks (ACN) using ARS.

Of particular interest are hand-held, cordless devices that do not require the intervention of a telephone company installer. A portable dialer offered by CES Industries (Model 335) is one

such device. It allows users to record and store up to twelve 16-digit numbers. The device is about the size of a package of cigarettes and is used by holding it directly over the receiver or by coupling it to the phone with a simple earphone-like extension. A similar device, Porta-Touch, will soon be offered by the BUSCOM Systems. A number of companies offer portable Touch-tone generators without memory. These include Soft-touch, another BUSCOM product, and Digitone, a product of Digitelle.

2) Costs. Memory dialers retail for about $125 each but can be bought in quantity for well under $100. Dialers without memory retail in the $30-$40 range.

2. Limitations of Telephone-Based Multiple Response Systems

Of the systems described within this category, the Percy VOXBOX is the only one which allows continuous processing of responses during the actual viewing of a program through a dedicated telephone line to the central computer. The operating and installation costs of the system are high, limiting the size and flexibility of the sample. While there are plans to expand it to other cities, it will not be available on a true multi-regional, let alone national, basis for some time. There thus remains a question about the external validity of responses made by 200 subjects in Seattle, Washington. There is also a question about the internal validity of a procedure which requires participants to actively respond (unlike the Nielsen Audimeter families who are unobtrusively monitored) over a long period of time. The system

also lacks flexibility in that all the responses relate to the standard eight-point adjective scale; additional questions could not easily be handled by the system. Furthermore, there are no specific quality controls to verify who is pushing the button.

PEAC is a highly sophisticated system which allows continuous monitoring of responses, although the in-home version is not currently available. The system will incur high installation costs, since a visit to each household would be required to install the unit and teach the individuals how to use it. Control could be maintained by requiring the respondent using the system to input a unique identification code. Although the in-home system could be implemented on a national basis, it is likely that it will be confined to a small number of metropolitan areas for some time to come.

Computafone is perhaps the most flexible of the various telephone-based, multiple response technologies. It can ask questions in a clear, natural voice, accept input from any kind of phone, and record open-ended responses. The automatic dialing capacity of the system could allow the system to conduct large random surveys with minimal human labor. However, the laws of many states require that the permission of the individuals must be obtained before the automatic dialer is utilized. Thus, the greatest potential benefit of this system goes for naught since respondent

recruitment costs still must be incurred. The system also has a relatively large input capacity. However, it is envisioned as a regional, rather than national, service that will be initially available in only a single market.

The major limitation of PAL, ARS and The Communicator is the need of Touch-tone telephones to input responses. It is estimated that only 38 percent of the telephones in the United States are Touch-tone and technical problems with telephone connections reduce that number slightly. However, there are Touch-tone adapters available which could easily be distributed to members of a sample so that they could respond. In fact, these devices might be more desirable than using the Touch-tone pad on the phone because they are portable and have memory. This would allow the respondent to record responses while watching a program and answer questions before going to the telephone. Also, the telephone number and respondent identification could be stored in the unit in order to facilitate access.

The Communicator is limited by the small number of responses (three) it can record per call. These limitations do not apply to PAL or to ARS. A problem unique to ARS is its rather garbled computer-synthesized voice. Unlike the other systems in this category, it could not be used to ask questions unless respondents were provided with a written copy in advance. However, ARS has the unique distinction of being the only audience feedback system currently in operation that can be accessed at no cost to the respondent from anywhere in the country.

3. Applications of Telephone-Based Multiple Response Systems

The limited capacity of these systems rules them out for most interactive programming applications. Conceivably, a sufficient number of phone lines could be attached to systems like PAL, ARS or The Communicator to handle the high peak loads of this application, but no system currently in use has sufficient capacity. Since Computafone is an outcall system (it places the call), it is not suited for interactive programming applications, where the respondent calls in.

The inability of the Percy System and the Communicator to identify respondents rules them out for public opinion polling. ARS, PEAC or PAL could be used for polling if respondents were recruited in advance and agreed to watch the program in process. Respondents could be identified by entering unique identification codes. However, PEAC and PAL will be geographically limited for some time to come. ARS cannot ask intelligible questions unless the respondents have the question protocol in front of them. All three systems require push-button phones for access, introducing a serious sample bias.

Computafone would be the ideal technology for polling, if legal prohibitions on "junk calling" were lifted. As it is, respondents have to be recruited and their permission secured in advance. The method of accomplishing this currently proposed by Computafone (i.e., mailing notifications in advance to respondents selected from reverse telephone directories) would probably

not be legal in some states. It certainly would not satisfy public opinion researchers, since only individuals with listed telephone numbers would be included in the sample.

For qualitative ratings, ARS is the only feasible system at present since it is the only one available on a truly national basis. However, it can only be used by owners of push-button telephones, introducing an unacceptable sample bias. It would be desirable only if used in conjunction with an automatic dialing device. The same strictures apply to PAL and The Communicator. The Percy System would not be ideal, even if nationally available, since it does not allow sufficient control over who is responding. Decision makers would have no way of knowing how different population segments react. PEAC II could be used in this application if it were to become available nationally. Its high storage capacity would make it especially attractive in this "data intensive" application. However, PEAC II, VOXBOX and ARS would all require a panel design to offset the costs of the terminals or automatic dialers.

Once again, Computafone is a definite possibility if an acceptable way can be found to work around the "junk calling" laws while not introducing unacceptable biases in the sample. If a panel study approach were adopted, it might be feasible to recruit subjects in advance by phone and then call them back at specific times during a "sweep" to collect their diary data. However, this

would be a rather tedious task for the respondents unless they had an off-line storage device. The alternative would be to draw a much larger sample and to perform telephone coincidentals focusing on qualitative responses during the sweep.

The ideal application for phone-based, multiple response systems is pretesting. Here, the sample biases associated with different techniques could be offset by the advantages of rapid turnaround and the relatively high data capacity of the systems. PEAC was designed specifically for such applications and, when available in the home, will have the unique capability to record both continuous ratings and answers to discrete questions with ease. The computer interface will make results available on an overnight basis. With the exception of VOXBOX, the other systems in this category do not have the capability of recording continuous ratings, unless such responses are first recorded on paper and then relayed manually to the computer. However, if they were used in conjunction with automatic telephone dialers, which could record responses off-line (much as PEAC would), this problem could be overcome. In that case, Computafone would be the most desirable since it could survey a large sample and perform any necessary follow-up calls with a high degree of efficiency. ARS and PAL are also possibilities, although they have the disadvantage of not having live operator intervention or automatic recording of open-ended responses as an integral part of the system. However, these capabilities could probably be added at a modest additional cost.

VOXBOX has limited utility as a pretest device despite its ability to collect continuous viewing data. The operators of the system are reluctant to use it in the invited viewing mode since theirs is primarily a syndicated data service. Specialized pretest studies invalidate the data they sell to their regular clients. Without invited viewing, the sample sizes are too small to serve the purposes of pretesting.

IV. A PILOT STUDY OF COMPUTERIZED PHONE-BASED DATA ACQUISITION

A common element across the interactive television technologies and their various applications is the process of collecting computer-readable data directly from the home though a direct interaction between the respondent and a computer accessed via the telephone. Since such technologies are not yet in common use, several questions arise concerning their feasibility. Will the public willingly and accurately use such systems? What are the potential advantages to the researcher? What barriers exist to widespread application of such systems? What hardware and software components must a viable system have?

To help answer these questions, we conducted a pilot study of computerized phone-based acquisition of data from the home. The technology examined was a phone-based, multiple response audience feedback system. In this test, users entered ratings into a computer using their own push-button telephones. The computer interface was the Consumers' Computer Corporation of America's PAL system (see Section III for a description).

The specific application tested was a qualitative ratings study. The approach was to simulate a single "sweep" of qualitative rating data collection employing computerized phone-based technology. The City of San Diego, California, the only city in which PAL was operational, was the test site. Viewing diaries were placed in randomly selected homes along with directions

for phoning in ratings to PAL. ACN asked to have the completed diaries returned to us so that we could compare the originals with the data output from PAL in order to assess the error rate and other patterns of use. Any difficulties in using the system were identified through inspection of the computer output, from follow-up contacts with users, and from a series of open-ended questions printed on the backs of the diaries.

Since the goal was to demonstrate data collection technology rather than the concept of qualitative ratings per se, we made a number of simplifying assumptions to reduce the "lead time" for the study and the magnitude of the tasks required of the respondents. Accordingly, ad hoc scales of enjoyment, worthwhileness and other qualitative dimensions were used without attempting to validate or pretest them in any formal sense.* Rather, the focus of pretesting was on the ability of respondents to understand the directions for phoning in their ratings. Instead of asking respondents to rate an entire week's programming, ten specific programs, including six PBS programs, were listed in the viewing diary. Respondents were also asked to rate two programs of their own choice; however, we required that only the ratings of a single program be phoned in to encourage participation by those who might have to leave their homes to gain access to a Touch-tone phone.

---

*The dimensions used were suggested by participants in the CPB technical conference on qualitative ratings (CPB, 1980a)

Two populations were of particular interest: regular PBS viewers and Hispanic viewers. Accordingly, two separate samples were recruited. A sample of KPBS, San Diego, subscribers was randomly selected from the station's mailing list. A second sample of Spanish-surnamed viewers was drawn from a reverse telephone directory of San Diego. A mail survey approach was used in an effort to minimize costs. However, this procedure yielded such a disappointing response rate among Hispanics that this aspect of the study was repeated using a combination of phone and mail recruiting techniques similar to those used by television ratings services to recruit minority group members for their diary studies.

In general, those who used the PAL system found it to be a convenient and enjoyable way of responding to questions about television programs. The few who found it to be excessively complicated or had a problem using it reacted more to special conditions that applied during the pilot test (e.g., redundant data entry by phone and by diary, calling the system before it was turned on, etc.) rather than to inherent flaws in the technology itself. Some potential advantages from the respondent's point of view were also revealed. One respondent enjoyed the longer time he had to consider his answers carefully, while another found the system less an invasion of privacy compared to conventional phone interview techniques. There was no evidence that users

considered a direct interaction with a computer to be either "dehumanizing" or an invasion of privacy.

Nonetheless, the overall response rate was quite low in comparison to conventional paper-and-pencil diary techniques. Once again, this was probably more a result of special procedures adopted for the pilot test than reflective of the technology itself. For example, recruiting for one phase of the study was conducted over a period of three weeks prior to its initiation. Those contacted in the last couple of days prior to the study exhibited response rates comparable to those observed using mail diaries, while those contacted earlier had much lower response rates. Language was another problem. Although all materials were translated into Spanish, the tone of the translation was evidently too formal and alienating for respondents whose dominant language was Spanish, and they responded in low numbers. Those who were bilingual or dominant in English responded at acceptable rates. The fact that the computer itself was English-speaking may also have contributed to the low response rates among Spanish-dominant subjects. Finally, it appears that a number of respondents believed that they were being asked to view all of the programs listed in the diary, despite repeated directions to the contrary. In studies that require viewing of a particular show, response rates are often as low as ten percent, somewhat lower than the rates observed here. Thus, the response rates were acceptable after allowing for some of the problems that were

unique to the pilot study. Additional attention to the problem of response rates, however, is clearly in order.

Data entry was remarkably error-free. Only 11 percent of all entries into PAL contained errors, using the paper-and-pencil diaries as the standard of accuracy. Since each transaction contained 24 characters, only about one-half of one percent of all characters entered were in error. About half of these would be detectable with interactive hardware sensitive to short character strings or illegal codes, so that the error rate theoretically could be greatly reduced. It should also be noted that a number of respondents who said that they used PAL did not actually have their responses recorded. These respondents apparently failed to enter the control character (#) that marked the end of the dialogue, or else hung up before PAL responded to the final entry, thereby cancelling the transaction.

Some of the potential strengths of PAL and similar phone-based computerized systems from the researcher's point of view were evident from the pilot study. Data were available immediately in computer-readable form, sparing the cost and time of manual data processing. Assuming a multi-wave study, the costs of the technology would soon be offset by the cost of mailing, printing and manual data processing of paper-and-pencil diaries. The progress of data collection could be monitored, and targeted follow-up phone calls could be made to correct problems that were evident from initial inspection of the data. Although not

implemented in the present study, such systems are capable of detecting errors in data input (e.g., short strings of numbers, illegal codes) and of requesting corrections as the data are entered.

Some potential barriers to widespread use of computerized phone-based systems were also apparent. Chief among these is the limited distribution of push-button telephones, estimated to be less than 40 percent nationwide. Respondents without push-button phones were asked to use public phones or phones in their place of work equipped with Touch-tone keyboards. Although a few respondents did use phones outside their homes, it is clear that they considered this a major imposition. This approach is not practical on a mass scale in studies requiring repeated data entry. Nor can all push-button phones be used to access the system. Some consumer-installed phones with push-button keyboards merely convert the input into rotary dial impulses, which cannot be detected by PAL. Other push-button phones have a condition known as "polarity reversal," stemming from improper installation, which means that Touch-tones are not registered by PAL after an initial connection is made. There are still many exchanges, especially in rural areas, that are not equipped to accept Touch-tones, whatever the nature and condition of the phone instrument. However, all systems can transmit Touch-tone <u>after</u> the initial connection is made.

Other difficulties were also encountered. The average transaction time was rather lengthy, as long as two and one-half minutes. This

could pose serious peak load problems for a nationwide system accessed by numerous telephone lines. However, there was evidence of a steep "learning curve"; frequent use of the system quickly resulted in average transaction times of one minute or less. There was some difficulty in coordinating the study with the day-to-day operation of PAL. Telephone lines were not always staffed by operators familiar with the study, so some user questions were not satisfactorily answered. Finally, an inherent limitation of the system is that only closed-ended questions can be asked. Open-ended comments cannot be recorded.

To make computerized phone-based systems practical for pretesting, qualitative ratings or interactive programs, several changes in the research procedures and system configuration used in the pilot test are in order. Most important is the addition of portable input devices (Section III) that would greatly expand the potential respondent pool beyond those who own Touch-tone phones. This measure would also reduce the average transaction time by as much as a factor of five, greatly relieving the peak load problem and reducing computer time charges.

A user dialogue with improved interactive capabilities should be developed that could detect and request correction on the most commonly made user errors. The system might even "walk" new users through questionnaires, although it would not be cost-effective to have the computer administer the questionnaire in regular use.

Procedures for interacting with respondents should also be upgraded. A live operator should be on standby to answer questions at all hours during which the system is in use. System software should be improved to generate real-time monitoring information that could be used to identify nonrespondents or those who make frequent errors so that follow-up calls could be immediately directed to them. An automatic phone dialer capable of delivering recorded messages could be used for these purposes.

The capacity of the system must also be increased. A number of WATS lines and additional "black boxes" to read the incoming Touch-tones, perhaps 15 to 20 in all, would be needed for a national service. A back-up system, perhaps employing telephone answering machines to record the Touch-tone impulses, would also be vital.

While it is easy to envision how to add a capability for recording open-ended comments to the system, it is not advisable to do so. The cost saving realized by eliminating manual data processing would soon evaporate while transaction times (and hence computer time changes) would skyrocket. A low-cost alternative would be to permit users to offer open-ended comments to the operator standing by to answer questions. Select respondents might be invited to make such calls based on their closed-ended responses to program content.

Finally, a parallel Spanish-speaking system should be implemented. Spanish-dominant respondents could be given special identification codes that would automatically route them to a Spanish-speaking software package when they place their calls.

V. THE COST OF AUDIENCE FEEDBACK SYSTEMS

A critical consideration in the application of interactive television technologies is their cost to the service user. While the technologies may offer distinct advantages over conventional data collection methods in terms of their functional capabilities, they will not be adopted by public broadcasters unless they offer a cost advantage as well. Choices between the various available technologies will also be made on the basis of their relative cost effectiveness. The goal of this section is to examine the issue of the cost effectiveness of the various interactive technologies both in comparison to conventional methods and in comparison with each other.

These estimates represent our best judgment of the possible costs. In virtually all cases, it is impossible to calculate cost figures with complete confidence and precision, since many of the systems have yet to be used in one or more of the applications considered in this paper. In other instances, the providers of the system have not as yet established a firm price structure for their services. This is not surprising since the actual costs will in large part depend on factors (e.g., response rates, phone line charges) that will not be quantifiable until the systems have been used a number of times in each application. The interactive technologies also introduce new dimensions to the cost equation for research. Computer connect times, terminal installation fees, terminal recovery rates, and amortization schedules for computer hardware are a few of the factors that

figure prominently but whose full impact on costs is unknown at present.

Therefore, for estimation purposes, a number of simplifying assumptions are necessary. We will limit our analysis to the prototypical examples described earlier for pretesting, qualitative ratings, public opinion polling and interactive programming. We will focus our analyses on the two or three most promising technologies for each application. With the exception of interactive programming, we will assume that the systems will be operated by research suppliers rather than by public broadcasting entities as an "in-house" function. We will use prevailing commercial rates for such component costs as sample recruitment, keypunching, computer time, live operators, etc. In so doing, we may err on the side of overestimating the actual costs, since an on-going research service should greatly reduce such costs by assuming the functions themselves, thereby realizing economies of scale. We will further assume a constant 50 percent overhead rate on the actual research costs. This rate is somewhat lower than that prevailing in many commercial research organizations. However, it is our assumption that one of the impacts of the new technology will be to reduce overhead by automating the research process and reducing recurring personnel overhead costs.

A. Cost Comparisons for Pretesting

In the commercial sector the cost of pretesting a half-hour program on a specialized target audience can range from as little

as $17 per head for an auditorium test to over $100 per head for focus group interviews. The most relevant comparison is with the service ASI provides in which cable viewers are invited to view a pilot and then are surveyed the following day via telephone. The cost of this type of study is about $20 per head. However, it is not exactly comparable to our prototype example since continuous ratings are not obtained. If this were done by conventional means (e.g., by having respondents record ratings in diaries), it would probably add $2 to $3 per head for manual data processing, postage, and follow-ups.

The cost effectiveness of audience feedback systems in pretesting depends on two factors: the nature of the transaction with the respondent and the number of test sessions for which respondents remain in the study. There are three different approaches to the respondent dialogue. First, the computer (e.g., Computafone or PAL) actually asks the questions and records replies made by push-button (PAL) and/or rotary dial (Computafone) impulses. A second approach is to have the respondents record their responses on paper in advance and enter them manually into the phone all at one time (preceding chapter). A third possibility is to enter the responses off-line into a portable terminal such as PEAC II or an automatic dialer) and then play them back into the phone in a short burst.

The first type of dialogue is probably not cost-effective in comparison with conventional techniques. In an hour's time, Computafone could probably complete no more than two such

transactions per telephone line, at a cost of $5.55 each for data collection. It would probably be less expensive to have a live operator collect the data and enter it directly into a computer, which is how many survey research firms presently operate. The potential cost benefits of Computafone in making the initial contact would be neutralized, since we are assuming that respondents would have been contacted in advance by mail or phone and would thus be equally available to either a live or computer operator when called back.

The same holds for the second type of dialogue in which users enter their written responses. However, here a machine could probably record a string of numbers with fewer errors more quickly than a live operator, so there might be a cost advantage to the automated system--perhaps as much as a 50 percent time savings. Either ARS or Computafone would be a possibility in this instance since the subjects would have the questionnaire in advance and would be able to follow the much simpler user dialogue despite ARS's heavily computer-accented "voice." If we assume that each dialogue would require the respondent to input a total of 60 numerical ratings and three short open-ended questions, each transaction might take five minutes for experienced respondents. In an hour's time, Computafone could complete perhaps 300 such calls at a cost of $1.67 per respondent. ARS would be slightly less expensive, with a total connect time of 30 hours, costing $600 for 400 calls, or $1.50 per head for

data collection. After adding the costs of data processing, recruiting and overhead, the per-head costs would be about $16 for either Computafone or ARS, compared to $18 for collecting the data with live operators in this fashion.

When data are collected off-line and played back in a burst, major cost savings could be realized over live operators. However, PEAC II might cost as much as $30 per head, due to its rather high installation and terminal costs. If a much less expensive and more portable terminal were used (perhaps by adding an inexpensive automatic telephone dialer to either ARS or Computafone), the costs could drop dramatically. Now, each transaction could take 30 seconds or less to complete. At that point, the distinction between the pricing policies of ARS and Computafone would probably tip the balance in favor of the former. Computafone charges by the hour, including the time it takes to place a call and get in touch with the designated respondent. This could add perhaps 90 seconds to the average call. With ARS, the clock starts running only <u>after</u> the call is placed. With ARS, the data collection costs would be only about 16 cents per head, compared with perhaps 40 cents to 50 cents per call for Computafone. Even after adding the cost of an additional WATS line and a live operator to the ARS cost (necessary to implement the collection of open-ended questions), the per-head cost for data collection would still be considerably less. APL could cost even less with its projected cost of 10 cents per transaction. However, this does not include the cost of adding WATS lines and additional computer hardware for the fast turnaround data analysis needed for pretesting.

Of course, the additional cost of the telephone dialer must be added to such a system. A related issue is the dialers' durability and the rate at which panelists will return them once a study is completed. Both factors are unknown at present. However, even if a relatively large number of terminals were destroyed or lost, say as many as 50 percent, such systems would still be cost-effective given a high enough number of repeated uses through the trade-off of computer connect time against terminal costs. Under the worst case (50 percent terminal mortalities), a repeat panel involving five tests would be necessary to equal the per-head cost of conventional data collection. Under the most favorable conditions, with up to 80 percent returned, the break even point would be as few as three repeat uses. If the terminal stayed in the home for as many as 20 tests, the per-head costs could drop to well under $10.

At this juncture, there are a number of crucial, unanswered questions bearing on the potential cost effectiveness of interactive technologies for pretesting. If terminals are used, the costs of replacement and maintenance are critical unknowns. As in other applications, there is also the question of the response rate that will be obtained using these technologies. A pilot study (Chapter IV) found low overall response rates, although among a population known for characteristically low rates of response. The true response rate will have a major bearing on the cost effectiveness of these technologies, especially in view of the fact that they must be embedded in a panel study to be cost-effective.

It is even possible that, in time, the response rates might be considerably higher than those observed with conventional methods. Efficient, real-time monitoring of incoming data may make highly targeted, highly cost-effective follow-up procedures possible that will boost return rates at a low marginal cost. Systems like PEAC II and ARS in which the respondent initiates the call may prove more convenient for respondents than conventional phone interviews.

It appears that Computafone or ARS could be cost-advantageous compared to conventional pretesting techniques if used in a mode where a panel of respondents enters a string of ratings without lengthy intervention by a live operator or without extensive prompting by the computer. However, only time will tell if the assumptions we have made about the response rates and productivity of these systems will be realized. If it proves feasible and acceptable to maintain panels over a relatively large number of tests, then the Audio Response Service or Computafone, used in conjunction with an automatic phone dialer memory, would have a distinct cost advantage. It might be possible to halve the cost of conducting pretests by conventional means.

PEAC II would not appear to have a cost advantage unless hardware and installation costs come down or unless an exceptionally large number of tests (e.g., once a week over a period of years) were conducted in each household.

B. Cost Comparisons for Qualitative Ratings

The cost of obtaining qualitative ratings by conventional means (i.e., diaries) is difficult to estimate, since only a single pilot study of the type of system envisioned in this report has been performed. A conventional viewing diary costs approximately $35 per head, which would certainly be the lowest possible cost for a qualitative rating diary. A recent field test of a qualitative rating diary suggests that actual costs may be two to three times this amount. For the sake of argument, we will use the figure of $60 per head per sweep for the prototype application described earlier in which a sample household would remain in the sample for four "sweeps" during a year.

Many of the cost considerations pertaining to pretesting apply equally to qualitative rating applications. In fact, for the present purposes, it is perhaps best to think of qualitative ratings as a kind of diagnostic testing in which respondents are empanelled for a relatively small number of repeat studies, but in which the volume of data generated by each study is considerably larger, e.g., 2,000 columns per sweep for a four-person household. Since the conventional means of data collection involve paper-and-pencil diaries, the trade-off shifts to comparisons of hardware and telephone communication costs with manual data processing and postage.

For example, for a family of four taking part in four qualitative rating "sweeps" during the course of a year, the cost of postage, printing and manual data processing could come to approximately $95 for the entire family. Alternatively, the same family

could receive nightly calls from Computafone during which each family member (or one family member serving as "secretary") would enter the rating data from the preceding day. Assuming that each family member watches seven programs a day and that ten columns would be needed to record each show, this would mean the entry of 280 columns of data per call. We estimate, based on our demonstration of the PAL system (Chapter IV), that this would take at least 15 minutes for the average person with a Touch-tone phone, and as much as a half-hour for a rotary dial phone. Assuming for the moment that respondents could be motivated to do this (a highly questionable assumption!), the average call, across rotary dial and Touch-tone, would cost about $5 in Computafone time, and collection charges would be at least $140 for the family of four. In addition, there would still be some minimal mail and printing charges involved, so that each family member could receive an individual diary. The Audio Response Service, with its higher per-minute charges, would be even more expensive if the data were entered manually in real time.

However, systems in which data are entered off-line offer some potential cost advantages. For example, if ARS were used in conjunction with an automatic phone dialer and if the dialer were shared by the entire family of four, the data entry printing and postage charges would combine to only about $20 for the

entire year for the family. This is because a relatively high volume of data can be played back quickly through the automatic dialers, as much as 200 characters in a minute-long phone call. In this application, family members would enter the data in the terminal as they watch the program, much as they would use a paper-and-pencil diary. Whether or not this approach would be cost-advantageous would mostly depend on the amortization of the dialer, its initial cost and the volume of data. Assuming an initial dialer cost of $80, it would be a break-even proposition when compared to paper-and-pencil diaries for a family of four. However, for a single person household, the break-even point would be a $40 dialer, since there would be less proportional savings in computer connect time relative to manual data processing of diaries. Another way to look at it would be that half of the terminals would have to be returned at the end of the year to make the system cost-effective in single person households. Assuming a high return rate (say 75 percent) and averaging across all households, we "guesstimate" a per-head cost of $45 per sweep.

PEAC II is another system that permits off-line data collection. The greater memory capacity of this system (500 columns vs. 192 for existing automatic telephone dialers) offers potentially greater cost savings in data transmission. However, the cost of the individual terminals and their installation is also much greater. Taking the initial PEAC II cost estimates at face value and assuming optimal utilization of the system, it would cost $30 per night for 16 nights over the course of the year,

or at least $480 to collect data from the family of four. In a single person household, the figure would be about $120. Averaging across households, this might amount to a per-head cost of as little as $30 per sweep. However, the implementation of a truly national qualitative rating service would introduce some major additional costs not covered in the initial estimates for PEAC II. For one thing, WATS lines would have to be added to the system. It is certain that the installation and training costs would rise dramatically with a national-level service, as opposed to one based in a single, confined geographical area. Moreover, the initial cost estimates are based on the assumption of a high (but unspecified) number of repeat uses in addition to qualitative rating sweeps. This might introduce an unacceptable bias for qualitative rating applications. If utilization had to be confined to qualitative ratings, the cost would be prohibitive.

Thus, electronic audience feedback systems do offer some major potential cost savings in qualitative rating applications. However, the word "potential" must be emphasized since at this point, there are a number of important, unanswered questions about the response rates, error rates, and hardware costs associated with these technologies.

C. Public Opinion Polling Cost Considerations

While we do not consider any of the interactive technologies suitable for public opinion polling at this time, a few comments

on their potential cost savings are perhaps in order. The most promising technology for such polling is the Computafone system. If the legal problems associated with "junk calling" could somehow be resolved, this system would permit highly cost-effective data collection. The system could possibly complete up to 200 ten-minute interviews per hour, at a cost (for data collection) of only $500. Even after adding the cost of instrument development, data analysis and interpretation and 50 percent overhead, the per-head cost could still be under $8. However, it is also possible that the completion rate would be less using an automated service than a live operator (e.g., "I don't talk to computers," click.) If so, the quality and cost-effectiveness of the results would suffer. This is an unknown parameter at this time.

D. Cost Comparisons for Interactive Programming

Interactive programming with mass call-ins has seldom been tried with conventional means. A somewhat analogous application is the telethon, a familiar phenomenon to public broadcasters. Public broadcasting telethons are done at exceptionally low costs since most stations maintain a large number of phone lines throughout the year and use volunteer labor to staff the phones. The volume of calls is probably considerably lower than would be observed during an interactive program. For comparison purposes, we will assume that interactive programs could be executed using a large number of dedicated phone lines (e.g., 75) and live operators who would record the "votes" of callers and manually tally them

so that results would be available within minutes. Further, we will assume that the live operators would be paid at the going rate for phone interviews, or $5 per hour, and would be employed for two hours per program or $750 per week. The 75 phone lines would cost $7,500 per month or $1,875 per week if used only once each week. We assume that the sample will be preselected through the use of prestamped, preaddressed post cards supplied by the audience at no cost to the station. Processing new applications would incur a minimal additional cost of perhaps $25 per week. After adding $200 per week for supervision and a ten percent mark-up, we arrive at a cost of $3,135 per show. If we assume that each "vote" will take 20 seconds to register, a bank of 75 phones could handle 4,000 calls in a 20-minute period at a cost of about 79 cents per head. The per-head costs would drop quickly with more frequent usage, making better use of the phone lines. If a program were represented on a daily basis, the costs would drop to as little as 25 cents per head.

If telephone answering machines were used instead of live operators, the cost would drop to 15 to 20 cents per head, assuming a relatively fast (one-year) amortization of the answering machines. Further savings would be realized if stations made use of the lines they maintain for fund raising. If the telephone line costs are eliminated in this fashion, per-head costs would dip to about 30 cents per head if live operators were used and to about 12 cents per head if a bank of answering machines were used.

With the parameters of our prototype example, AT&T's Dial-It service would cost about 17 cents per head to the station. Of course, it would cost the viewer 50 cents for each use, and it might not be politically acceptable (i.e, "elitist") to pass this cost on to the viewer. The principal costs to a local station for a weekly call-in show would be one-fiftieth of the $5,000 annual subscription fee, $500 for an hour's use of the system in a single time zone, a $50 surcharge for minute-by-minute results, and about $75 per week for overhead and processing applications from users. The cost to the local station could be reduced somewhat if an entity such as CPB were to absorb the $5,000 yearly fee and reallocate it to a number of individual stations.

If a local station were to use the VOTRAK system for a weekly call-in program, the cost per use would be one-quarter of the $6,000 monthly lease plus $500 dollars per week for telephone lines charges. Adding overhead and administrative costs would bring the total weekly costs to $2,227, or about 56 cents per head. Full utilization (i.e., spread across ten shows per week) could dramatically reduce costs to as little as 28 cents per head. Alternatively, a station might make "free" use of the lines it keeps on hand for fund raising, in which case the per-head cost would drop to about the same as Dial-It's if the system were fully utilized. However, if VOTRAK were to be shared by a number of stations in different parts of the country, WATS lines would have to be used instead of local lines making the total cost prohibitive. What's more, we are not entirely confident that the 20 lines recommended would be sufficient.

Twenty lines could handle 4,000 calls in a 20-minute period only if the lines were used with peak efficiency and if each transaction took no more than six seconds to complete. Even with preselection and scheduling of the callers, momentary peak loads would be bound to exceed the capacity of 20 lines. Ten additional lines would raise the per-head cost of a weekly show to 84 cents or to 61 cents if "free" phone lines already owned by the station were used.

The costs of some interactive systems (e.g., VOXBOX, PEAC II) would be clearly prohibitive in this application due to either excessive lines costs or hardware costs. Since the sample would represent only the viewers of a particular program, it would not be feasible to use the sample for other purposes (e.g., pretesting or qualitative ratings). In fact, such a system would only be cost-effective if it were part of a multi-purpose, two-way network normally dedicated to applications other than research (e.g., banking, shopping by phone).

Although not ideally suited for the interactive programming applications, systems like PAL and the Audio Response Service could offer a distinct cost advantage. If the Audio Response Service were used and if each transaction lasted only 15 seconds, roughly 17 hours of connect time would be required to process 4,000 calls at a cost of only $340. Assuming modest additional costs for data processing and administration, the cost would be only about 11 cents per head. This assumes that SBC would be willing

to add extra lines to their system at little or no additional hourly cost. However, even if the hourly cost were doubled, the per-head cost would rise to only about 18 cents per head. Of course, such a system could be used only by owners of push-button telephones.

Thus, in terms of cost to the station, Dial-It has a clear cost advantage over VOTRAK unless a station's own "pledge lines" were used to reduce the cost of the latter. The adaptation of the Audio Response Service or a similar service could substantially reduce the per-head costs. However, we must also note that while cumbersome, a bank of telephone answering machines (or even a battalion of live operators!) might be more cost-effective than any of the technologies best suited for interactive programming applications. This is especially true if the station could make free use of the multiple telephone lines it maintains for annual fund raising.

A final consideration is the front-end cost -- the size of the initial investment a station would have to make to "buy into" interactive television. Of the two leading technologies, Dial-It has a slightly lower initial cost ($5,000 vs. $6,000 for VOTRAK). The Audio Response Service would have a very modest initial outlay -- the $100 minimum charge for a single month's use. Most expensive would be the purchase of a bank of answer phones. An initial investment well in excess of $10,000 would be required.

VI. CONCLUSIONS AND RECOMMENDATIONS

No single audience feedback system meets the operational requirements of pretesting, qualitative ratings, public opinion polling and interactive programming. However, we can suggest alternatives for each application which offer distinct advantages over conventional methods and their competitors in terms of functional capabilities and cost. We must caution the reader that these recommendations are based primarily on information provided by the suppliers and marketers of the technologies. Operational evaluations of these systems were not performed in all cases.

A. Conclusions

Our conclusion is that interactive television is now a reality. It can be placed in the service of public broadcasters throughout the country. The advent of widely distributed, broadband communications networks -- still years away -- is not a necessary prerequisite.

Audience feedback systems offer some potential, significant advantages to public broadcasters, particularly in the areas of pretesting and interactive programming. The various interactive technologies could allow public broadcasters to put themselves in touch with their audiences in ways which were heretofore impossible, at a cost that public broadcasters can afford.

Significant cost advantages might also be realized in qualitative rating applications. However, "instant" public opinion polling is not feasible at this time. The technologies which have the greatest promise for each application are discussed below.

1. Prestesting

The use of audience feedback systems for pretesting television pilots appears most attractive only when viewer responses can be "batch loaded" directly into a computer without live operator intervention or extensive computer promoting. Computafone is potentially superior to conventional methods and its "new technology" competitors both in terms of its operational capabilities and cost. The system allows direct input of responses from any home phone into a computer. It also has the capability of recording open-ended comments. Perhaps the greatest cost savings are realized by the automated call-back capability the system possesses. However, the impact of this capability is limited by laws against automatic calling, which require that respondents be initially contacted by mail or live phone interviews. The Service Bureau Corporation's Audio Response Service is another alternative, although live operators would have to be added to the existing configuration to permit collection of open-ended responses, and it is limited to homes with Touch-tone phones. ARS is the only system suitable for pretesting that is currently available to local stations across the nation.

Both Computafone and ARS become especially attractive when used in conjunction with simple off-line storage devices. In a long-term panel study, the costs of pretesting could fall to under $10 per head, less than half the current rate for conventional methods. Either system offers data collection from the natural viewing environment with unprecedented speed. Although less attractive in terms of cost, PEAC II is the one technology on the horizon which can most readily collect continuous responses to a program in progress.

2. Qualitative Ratings

As with pretesting, substantial cost savings might be realized by using the new technologies over conventional (diary) techniques through reductions in labor, postage and manual data processing. However, this savings would only apply if telephone-based multiple response systems such as Computafone or ARS are used in conjunction with off-line storage devices. PEAC II would have a relative advantage in the amount of data it can store, but would be cost-advantageous only if the costs of home terminal installation and user training could be amortized over a large number of other (e.g., pretesting, marketing) studies.

3. Public Opinion Polling

Primarily because of their failure to meet the rigorous sampling requirements of public opinion polling, we do not believe any of the existing technologies could replace live telephone interviewers for this purpose. The possible exception is Computafone,

should the legal problems relating to "junk calling" be resolved.

4. Interactive Programming

If restricted to audience participation formats (as opposed to those requiring a representative sample), audience feedback systems have much to offer in interactive programming applications. The most cost-effective alternative of all might be a bank of telephone answering devices attached to incoming telephone lines, especially if the telephone lines are those which a station maintains for pledge and auction purposes. However, this would be a rather cumbersome procedure, and it requires a prohibitively large capital outlay to purchase the machines. AT&T's Dial-It service is also a possible alternative, though the biggest drawback to its usage is the 50 cent charge to each caller. In instances in which stations maintain a large number of telephone lines, VOTRAK becomes a feasible option.

5. User Acceptance

A major uncertainty underlying all the prototype applications discussed here is the level of respondent acceptance of electronic audience feedback systems. Our limited experience indicates that respondents are willing to participate in research activities utilizing these technologies and use them with a high degree of accuracy. However, additional research must be undertaken to further analyze this question. Investigations of methods of increasing response rates, including real-time monitoring of incoming data and highly targeted follow-up activities, are a top priority.

B. Recommendations

In broad perspective, we recommend that public broadcasting in general and the Corporation for Public Broadcasting in particular move to take advantage of the unique potential offered by emerging audience feedback technologies. At the same time, we must stress the word "emerging," cautioning us that the time is right for further exploration of the various systems, rather than a major commitment to any one of them for the present. We believe that CPB can play an important role in supporting the application of the various technologies to the specific needs of local stations and educating them as potential users.

We recommend the following specific steps:

1. Continue to monitor the development of the more promising technologies described in this report. A simple first step would be for CPB to follow-up on our initial contacts, placing themselves and interested local stations on the mailing lists of services like Computafone, ARS, and PEAC.

2. Subsidize further demonstration projects, such as our PAL study, involving local stations and local research providers which would apply services (e.g., PAL, PEAC II, Computafone, QUBE) that are currently available only on a local basis but that may someday be available nationally. For comparison purposes, the demonstrations should be designed to meet common specifications, such as the ones contained in the "prototype applications" in Section II.

3. Mount an educational campaign designed to promote the use of the two systems currently readily available on a national basis: Dial-It (for interactive programming) and ARS (for pretesting). CPB might support these applications by such actions as absorbing Dial-It's $5,000 annual fee or by subsidizing the design of standard questionnaires and sampling procedures for ARS.

4. Since the use of simple off-line data storage devices is critical to the cost effectiveness of many systems across uses, CPB should fund a pilot test of these devices in either a pretesting or qualitative rating application. The pilot test should provide information about user acceptance, response rates and terminal return rates.

5. Interactive programming is a fundamentally new concept for which the key to acceptance will be the creative opportunities perceived by the programmer rather than the cost effectiveness or relative advantage of the technology itself. To promote interactive television as a creative "plaything," it might be helpful to encourage its exploration by a single station across a wide variety of programs, perhaps in conjunction with the Program Fund. In this case, VOTRAK, rather than Dial-It, would be the preferred alternative based on cost effectiveness if the station's unused pledge lines were used.

REFERENCES

Cablevision. "The Dilemma on Cable's Two-Way Street: Is Now the Right Time to Go With Interactive, Addressable Technology?" Cablevision, March 2, 1981, pp. 26-29.

Chen, M. et al., 3-2-1 Contact Test Show Evaluation. (New York: Children's Television Workshop, 1979).

CPB. Proceedings of the 1980 Technical Conference on Qualitative Television Ratings. (Washington, D.C.: Corporation for Public Broadcasting, 1980a).

CPB. Proceedings of the 1980 Applications Conference on Qualitative Television Ratings. (Washington, D.C.: Corporation for Public Broadcasting, 1980b).

Los Angeles Times. "Privacy and 2-Way Cable TV." Los Angeles Times, March 3, 1981, Section 6, pp. 1, 7.

Media Science Newsletter. "Use of Telephone Two-Way Increases Audience Size." Media Science Newsletter, October 16-31, 1980, p. 3.

APPENDIX

AUDIENCE FEEDBACK SYSTEMS

Telephone-Based Single Response Systems

Dial-It	American Telephone & Telegraph Company (contact local telephone company)
VOTRAK	Unlimited Television, Inc. 324 East 35th Street New York, NY 10016 (212) 725-5546

Telephone-Based Multiple Response Systems

VOXBOX	R.D. Percy & Company 3712 Bank of California Center 900 Fourth Avenue Seattle, WA 98164 (206) 622-4755
PEAC	PEAC Developments 76 Ferris Road Toronto Ontario M4B 164 CANADA (416) 968-3679
Computafone	Media Service Measurements 324 East 35th Street New York, NY 10016 (212) 725-5546
The Communicator	Information Technology 4955 E. Anderson Fresno, CA 93727 (209) 255-8600
Audio Response Service	The Service Bureau Company 500 West New England Avenue Greenwich, CT 06830
PAL	Consumers' Computer Corp. of America 6815 Convoy Court San Diego, CA 92111 (714) 571-5610

CES	Communications Electronics Specialities, Inc. 260 West New England Avenue Winter Park, FL 32789 (305) 645-0474
Soft Touch Porta Touch	Buscom Systems, Inc. 4700 Patrick Henry Drive Santa Clara, CA 95050 (800) 538-8086
Digitelle	Digitelle 21 Dixon Avenue Copiague, NY 11726 (516) 842-8885